CUSTOMER NO. 30430

PATENT APPLICATION
Docket No. 02-HK-076

CERTIFICATE OF MAILING BY EXPRESS MAIL
"EXPRESS MAIL" Mailing Label No.

Date of Deposit: 07-25-03.
I hereby certify that this paper or fee is being deposited with the U.S. Postal Service "Express Mail Post Office to Addressee" on the date indicated above and is addressed to Mail Stop Patent Application, Commissioner for Patents, Post Office Box 1450, Alexandria, VA 22313-1450

Signature:
Name: Marco Barbarash

PACKET BUFFER CIRCUIT AND METHOD

BACKGROUND OF THE INVENTION

Technical Field of the Invention

The present invention relates to a buffer for storing information, such as packetized data.

Description of the Related Art

Driven by growing bandwidth demands from the ever-increasing population of network users, there is a relatively large market for various kinds of networking equipment. Among them, packet networking systems such as packet switches and routers are the key building blocks of networking infrastructures.

Generally, a packet switch/router performs two major functions: packet routing and forwarding. The former looks up the route-table to decide where an incoming packet will be forwarded (i.e., by which output port of the switch/router that the packet will leave); and the latter executes the actual forwarding operations. Before a packet can be forwarded to the next hop, due to resource contentions or other reasons, it may need to be stored in a packet buffer somewhere in the switch/router. In order to absorb temporary traffic congestions, fast and dense packet buffers are indispensable for the building of a high-performance fast packet switches/routers. Further, intensive research on high-speed switches/routers in the past decade have revealed that packet buffers used in a switch/router constitute a bottleneck for cost reducing and performance improving. Up to now, it still remains a difficult challenge to make fast and dense packet buffers which meet the needs of high-speed networking applications.

Packet buffers are generally solid-state random access memories (RAM) built by CMOS technologies. Generally speaking, those CMOS RAMs can be classified into two major

categories: static random access memory (SRAM) and dynamic random access memory (DRAM). The former is faster and the latter is denser. On one hand, one can build a fast and small buffer using SRAM; on the other hand, one can build a slow and dense buffer using DRAM. However, neither pure SRAM nor pure DRAM can build a desired fast and dense packet buffer.

A two-hierarchy SRAM+DRAM architecture was previously proposed in the art to build a first-in-first out (FIFO) fast and dense packet buffer. In this architecture, a DRAM provides the main storage capacity and a small SRAM is located between the external access interface and the DRAM to serve as a cache for access acceleration. Further, special properties held by FIFO packets are utilized to pipeline the operations of SRAM and DRAM. While this solution may represent a right direction for attacking the intended problem, nevertheless, it requires a sophisticated scheduler and limits its applicability to FIFO buffers only, which offsets the benefits obtained from this solution.

Based upon the foregoing, there is a need for a packet buffer which is sufficiently sized to hold a relatively large

number of packets, sufficiently fast to accommodate relatively high speed communication and relatively simple in implementation.

SUMMARY OF THE INVENTION

Embodiments of the present invention overcome shortcomings associated with prior packet buffers and satisfy a significant need for a relatively fast and dense buffer for storing packets of data.

According to an exemplary embodiment of the present invention, there is provided a buffer including a first memory and a second memory, the second memory being less dense but faster than the first memory. At least one queue is associated with the first and second memories and serves as a pointer to point to the locations in the first and second memories where data packets are stored. Incoming data packets are initially stored in the first memory. Individual data packets are subsequently transferred from the first memory to the second memory upon the individual data packet becoming the head-of-line packet, i.e., the packet appearing at the top of the at least one queue. By maintaining the

head-of-line packets in the fast second memory while maintaining the remaining packets in the denser first memory, the buffer is advantageously capable of providing better performance in a relatively efficient manner.

A method of operation may include storing incoming packets of data in a first memory, the packets being associated with the at least one queue, and transferring a first packet of the incoming packets from the first memory to the second memory upon the first packet becoming the head-of-line packet for the at least one queue. The first packet is sent to a telecommunications device following the first packet being stored in the second memory.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the system and method of the present invention may be obtained by reference to the following Detailed Description when taken in conjunction with the accompanying Drawings wherein:

Figure 1 is a block diagram of a packet buffer according to an exemplary embodiment of the present invention;

Figure 2 is a flow chart illustrating an operation of the packet buffer of Figure 1;

Figure 3 is a block diagram of a system having the packet buffer of Figure 1 therein; and

Figure 4 is a flow diagram for forwarding a stored packet.

DETAILED DESCRIPTION OF THE EXEMPLARY EMBODIMENTS

The present invention will now be described more fully hereinafter with reference to the accompanying drawings in which exemplary embodiments of the invention are shown. This invention may, however, be embodied in many different forms and should not be construed as being limited to the embodiments set forth herein. Rather, the embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art.

Referring to Figure 1, there is shown a packet buffer 1 according to an exemplary embodiment of the present invention. Packet buffer 1 is adapted for relatively high-speed applications, such as use in telecommunications networks. Packet buffer 1 may include a first memory 3 and

a second memory 5 that is faster but less dense than first memory 3. Packet buffer 1 may include a control buffer 7 for pointing to locations in first memory 3 and second memory 5 where individual data packets are stored. In general terms, packet buffer 1 transfers data packets between first memory 3 and second memory 5 based upon the status of control buffer 7.

As stated above, first memory 3 is larger than second memory 5 but has slower access times than second memory 5. According to the exemplary embodiment of the present invention, first memory 3 is a dynamic random memory (DRAM). In order to provide even faster read access times, first memory 3 may be a destructive-read DRAM (drDRAM). In other words, first memory 3 may be a DRAM in which a piece of data read from the DRAM is destroyed and no longer maintained in the DRAM. In a destructive-read DRAM, a write-back cycle does not follow a read cycle so as to restore the data read from the drDRAM. First memory 3, in this case, operates as a read-once memory. Because write-back cycles are not performed in the drDRAM, first memory 3 provides noticeably

faster access times, relative to conventional DRAMs which perform write-back cycles.

In choosing the particular configuration for first memory 3, a pair of competing interests may be considered. On one hand, there is a desire to configure first memory 3 as having relatively fewer rows of memory cells, each of which has a relatively large number of memory cells. Because first memory 3 is a dynamic memory, it must be periodically refreshed, so fewer rows of memory cells advantageously result in the total refresh time for refreshing first memory 3 being reduced. Consequently, the amount of time available for accessing first memory 3 increases.

On the other hand, it is advantageous for reasons of data storage efficiency to have a relatively large number of rows, each of which contains a relatively few number of memory cells. Consider a packet stored in consecutive memory cells of first memory 3. If even a single memory cell in a row is used for storing a packet, the contents of the remaining memory cells in the row will be lost when the packet is read from first memory 3, due to first memory 3 being a drDRAM. In order to avoid the loss of packet data,

first memory 3 may be utilized such that each row of memory cells may only store information corresponding to a single packet of information. However, this utilization of first memory 3 disadvantageously results in memory cells being unused. To address the need for a drDRAM having a relatively small number of relatively large memory cell rows as well as the need for a drDRAM having a relatively large number of relatively small rows, first memory 3 may be configured as follows.

According to an exemplary embodiment of the present invention, first memory 3 may utilize the concept of logical rows when accessing memory cells. Specifically, a physical row in first memory 3 may be divided into two or more "logical" rows, each of which is individually addressed. In other words, a logical row may be accessed for performing a read operation without accessing the other logical row(s) in the same physical row as the accessed logical row. In this way, only an accessed logical row(s) in a physical row will lose its data following a memory read operation to the accessed logical row. By storing information from different packets in different logical rows in the same physical row,

the utilization of logical rows in first memory 3 for storing and retrieving packets allows for the physical rows of first memory 3 to be relatively large. Consequently, data storage efficiency is increased while total refresh time is decreased.

The utilization of logical rows may be implemented using a two level addressing hierarchy. For example, first memory 3 may have a multi-divided word line architecture wherein each logical row in a physical row has a distinct word line. In this way, each logical row is individually addressable, so information from more than one packet may be stored in the same physical row of memory cells. Further, the word lines for addressing memory cells in each physical row also may be enabled at the same time. This allows for multiple logical rows in the same physical row to store information from the same packet, and also allows for a refresh operation to advantageously refresh an entire physical row of memory cells at the same time.

Second memory 5 is adapted to provide faster access times than the access times provided by first memory 3. In accordance with an exemplary embodiment of the present

invention, second memory 5 may be a conventional static random access memory (SRAM) that maintains its stored data values following access thereof. Alternatively, second memory 5 may be another type of memory that maintains stored data values following the stored data being read therefrom.

Control buffer 7 may include one or more queues 9. For each queue 9, each entry of the queue serves as a pointer to a physical address location in first memory 3 or second memory 5 where a corresponding data packet is stored. Each queue 9 may perform as a first-in-first-out (FIFO) memory in which the order of data read from the FIFO follows the order of data written thereto. In the event control buffer 7 includes two or more queues 9, each queue 9 may store data packets corresponding to a certain characteristic. For example, data packets may be organized in queues 9 according to packet destination or Quality of Service (QoS) class.

The one or more queues 9 may be implemented using at least one SRAM, as is known in the art. Control buffer 7 may further include control circuitry 11 which receives packets, generates corresponding address values for the data packet and determines the particular queue 9 into which the address

values are to be stored. Further, control circuitry 11 may generate necessary control signals (CE, RD/WRB, etc.) provided to first memory 3 and second memory 5 for storing received packets therein, i.e., for causing first memory 3 and second memory 5 to selectively perform memory access operations. Control circuitry 11 may also control the transfer of packets from first memory 3 to second memory 5, as will be described in greater detail below. Control circuitry 11 may be implemented with timing and/or logic circuitry for controlling queues 9, first memory 3 and second memory 5.

It is understood that circuitry other than control circuitry 11 may be used to provide control signals for controlling first memory 3 and second memory 5 in order to store received packets.

Figure 2 illustrates a system or network 20 in which packet buffer 1 may be utilized. System 20 may include a plurality of nodes 22, each of which is capable of communicating packets of data to each other and to other devices located in the system or elsewhere. Each node 22 may include at least one packet buffer 1, which is coupled to at

least one communications port 24 through which packets of information may be transmitted and received.

The operation for receiving a packet in accordance with an exemplary embodiment of the present invention will be described with reference to Figure 3. It is understood that embodiments of the present invention are not necessarily limited to the particular order of steps described below and illustrated in Figure 3. Rather, the order of steps may be reasonably modified from that described and illustrated.

In the event that a packet received at a node 22 cannot be immediately forwarded to another node, the packet may be received by packet buffer 1 for temporary storage therein. Initially, control circuitry 11 or other control or decode circuitry (not shown) in packet buffer 1 identifies the particular queue 9 to which the received packet is to be associated. Queue identification may be performed, for example, by examining the header of the packet associated with the received packet. Next, a determination is made, such as by control circuitry 11, as to whether the identified queue 9 is empty or instead contains at least the address of the HOL packet. In the event the identified queue 9 is

empty, the received packet is stored in second memory 5 (SRAM) and the address of the received packet is stored in the identified queue 9. This ensures that the HOL packet of the identified queue 9 is available from second memory 5 when the HOL packet is to be delivered from packet buffer 1 to the next communications node 22.

However, if the identified queue 9 is not empty, a determination is made, such as by control circuitry 11, as to the type of the received packet. In the event the received packet is a fanout splitting multicast packet and therefore capable of being forwarded to a number of destinations individually at separate times, the received packet is stored in second memory 5 (SRAM). By being stored in second memory 5, the received packet is advantageously available to be read from second memory 5 a number of times (for multiple transmissions) without being corrupted.

If the identified queue 9 is not empty and if the received packet is a nonfanout splitting multicast packet or a unicast packet and therefore be capable of being forwarded to a destination node at a single time, the received packet is stored in first memory 3 (drDRAM).

Figure 4 illustrates an operation for forwarding a stored packet to a destination node, in accordance with an exemplary embodiment of the present invention. It is understood that embodiments of the present invention are not necessarily limited to the particular order of steps described below and illustrated in Figure 4. Rather, the order of steps may be reasonably modified from that described and illustrated.

In this example, a series of packets associated with a particular queue 9 is sequentially retrieved from packet buffer 1 for subsequent transmission to the predetermined destination node. Specifically, the address of the HOL packet is provided by the particular queue 9. With the address of the HOL packet, the HOL packet is retrieved from second memory 5 (SRAM) and placed on the data I/O for subsequent transmission to the destination node. The particular queue 9 is updated to point to the new HOL packet. In the event the particular queue 9 is not empty following it being updated, the address of the new HOL packet may be retrieved from first memory 3 (drDRAM) and placed in second memory 5 (SRAM). Further, the new address of the new HOL

packet is placed in the particular queue 9 so that the particular queue is able to point to the correct location in second memory 5 where the new HOL packet is stored. At this point, packet buffer 1 is ready to provide another packet for subsequent transmission to a destination node.

The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as a departure from the spirit and scope of the invention, and all such modifications as would be obvious to one skilled in the art are intended to be included within the scope of the following claims.